Institut de Recerca en Economia Aplicada Regional i Pública
Research Institute of Applied Economics

Document de Treball 2020/13, 42 pàg.
Working Paper 2020/13, 42 pag.

Grup de Recerca Anàlisi Quantitativa Regional
Regional Quantitative Analysis Research Group

Document de Treball 2020/05, 42 pàg.
Working Paper 2020/05, 42 pag.

"In knowledge we trust: learning-by-interacting and the productivity of inventors"

Matteo Tubiana, Ernest Miguelez and Rosina Moreno

UBIREA
Research Institute of Applied Economics

AQR
Regional Quantitative Analysis Research Group

Institut de Recerca en Economia Aplicada Regional i Pública
UNIVERSITAT DE BARCELONA
WEBSITE: www.ub-irea.com • CONTACT: irea@ub.edu

AQR
Grup de Recerca Anàlisi Quantitativa Regional
Regional Quantitative Analysis Research Group
WEBSITE: www.ub.edu/aqr/ • CONTACT: aqr@ub.edu

Universitat de Barcelona
Av. Diagonal, 690 • 08034 Barcelona

The Research Institute of Applied Economics (IREA) in Barcelona was founded in 2005, as a research institute in applied economics. Three consolidated research groups make up the institute: AQR, RISK and GiM, and a large number of members are involved in the Institute. IREA focuses on four priority lines of investigation: (i) the quantitative study of regional and urban economic activity and analysis of regional and local economic policies, (ii) study of public economic activity in markets, particularly in the fields of empirical evaluation of privatization, the regulation and competition in the markets of public services using state of industrial economy, (iii) risk analysis in finance and insurance, and (iv) the development of micro and macro econometrics applied for the analysis of economic activity, particularly for quantitative evaluation of public policies.

Abstract

Innovation rarely happens through the actions of a single person. Innovators source their ideas while interacting with their peers, at different levels and with different intensities. In this paper, we exploit a dataset of disambiguated inventors in European cities to assess the influence of their interactions with co-workers, organizations' colleagues, and geographically co-located peers, to understand if the different levels of interaction influence their productivity. Following inventors' productivity over time and adding a large number of fixed effects to control for unobserved heterogeneity, we uncover critical facts, such as the importance of city knowledge stocks for inventors' productivity, with firm knowledge stocks and network knowledge stocks being of smaller importance. However, when the complexity and quality of knowledge is accounted for, the picture changes upside down and closer interactions (individuals' co-workers and firms' colleagues) become way more important.

JEL Classification: O18, O31, O33, O52, R12.

Keywords: Inventors, Productivity, Stock of knowledge, Interactions.

Matteo Tubiana: University of Bergamo. Email: matteo.tubiana@unibg.it

Ernest Miguelez: GREThA UMR CNRS 5113 - Université de Bordeaux. Email: ernest.miguelez@u-bordeaux.fr

Rosina Moreno: AQR-IREA Research Group, University of Barcelona. Department of Econometrics, Statistics and Applied Economics. Av. Diagonal 690, 08034 Barcelona, Spain. Tel. (+34) 934 021 823. Email: rmoreno@ub.edu

Acknowledgements

The authors acknowledge constructive comments received from Cristiano Antonelli, as well as participants at the 4th Geography of Innovation conference, Barcelona 2018; the Technology Transfer Society Conference, Valencia 2018; the 5th Global Conference on Economic Geography, Cologne 2018; the XLV Reunión de Estudios Regionales, 2019; and the Creative Sparks Workshop, Berlin 2019. Ernest Miguelez and Rosina Moreno acknowledge the financial support provided by the Ministerio de Economía y Competitividad for the project entitled 'Innovation and locational factors: Diversification, knowledge and the environmental revolution', ECO2017-86976-R, and Ernest Miguelez from the French National Research Agency (TKC project – reference: ANR-17-CE26-0016). All errors remain our own.

Introduction

Knowledge diffusion is critical for innovation and an engine of economic prosperity (Lucas 1988; Romer 1990). However, knowledge does not readily diffuse from person to person, and from organization to organization. It requires effort to be shared and tends to evolve and transform as actors receive and construct upon it. Individuals' learning and their subsequent productivity rely, first, on their efforts and capabilities – their human capital, as well as with whom they interact (Lucas 2009; Lucas and Moll 2014; Akcigit et al. 2018). Innovators, indeed, recombine new pieces of knowledge they learn while interacting with others (Weitzman 1998).[1] Nevertheless, learning is bounded within the social and geographic limits of individuals' interactions (Boschma, Balland, and Kogler 2015) and by the level of trust of the established relationships (Akçomak and ter Weel 2009; Miguélez, Moreno, and Artís 2011). When trust is high, information moves smoothly and facilitates interactive learning (Boschma 1999; Maskell and Malmberg 1999).

In this framework, this study aims at understanding the knowledge dynamics of individual inventors, a representative group of knowledge workers. In particular, we put forward a framework to investigate inventors' opportunities to interact with their peers' knowledge stocks at the territorial (city), organizational (firm) and personal (individual's network) levels, and their subsequent capacity to produce more and better ideas.

The choice of these three layers is grounded in well-established strands of literature. It is well understood that knowledge is regarded to be sticky in space and resists diffusion beyond too far geographic limits. As the relevant ingredient of knowledge is tacit, it is hard to exchange it in markets and knowledge externalities occur among geographically close actors (Boschma and Frenken 2010). Therefore, cities are seen as platforms to reduce the costs of interacting and allow accessing spatially sticky knowledge (Duranton and Puga 2001; Storper and Venables 2004; Carlino and Kerr 2014).

[1] To the best of our knowledge, the concept of "learning by interacting" was introduced in innovation studies originally by Lundvall (1988) and referred to how interaction between producers and users in innovation enhances the competence of both. It has been widely adopted by the "regional innovation systems" and the "learning region" literatures to depict how knowledge is to be diffused locally (Lundvall 1996; Lambooy 1997; Boschma 1999; Andersson and Karlsson 2002; Andersson and Karlsson 2004) – see also Nooteboom (2000), for a more managerial perspective.

However, researchers have lately questioned the idea that knowledge is "in the air" (Marshall 1890) in cities and clusters. In contrast, knowledge diffuses through planned and well-structured alliances between individuals, firms and other organizations (Fitjar and Rodríguez-Pose 2017). In parallel to these developments, the knowledge-based view of the firm has explored the firm's role in organizing and distributing knowledge and seeking competitive advantage internally (Kogut and Zander 1992). This literature contents that the mere existence of the firm facilitates the transferability of knowledge between individuals within firms' boundaries (Allen 1977; Teece 1977; Grant 1996).

The contributions of this paper are manifold. First, we extend the research strand about knowledge generation at the individual level, which is increasingly considered as the fundamental level of analysis for exploring knowledge creation mechanisms (Fleming 2001), but still under-investigated – except for few papers such as Giuri and Mariani (2013), Akcigit et al. (2018), Jaravel, Petkova, and Bell (2018) and Moretti (2019). Specifically, we concentrate on inventors, a class of highly skilled, highly educated knowledge workers who are behind the production of technological innovations.

Second, we assess the importance for inventors' productivity of accessing knowledge from three different layers: the city, the firm, and the inventor network. While we expect the impact of the three levels to be positive and significant, we are agnostic on which layer is going to dominate. Extant research generally focuses on only one level of analysis while neglecting the others, with the empirical consequences of ignoring unobserved heterogeneity coming from the other layers.

Third, we focus on European cities or metropolitan areas in the territorial analysis.[2] There are substantial theory and evidence that innovation is primarily an urban phenomenon (Carlino and Kerr 2015). By contrast, most empirical studies on the geography of innovation have used administrative boundaries such as NUTS2 regions in Europe or the US States. Our chosen spatial scale of analysis should reflect more closely the dynamics of knowledge interactions and innovation, and complement the existing empirical evidence at the level of regions.

---

[2] Even though our unit of analysis is officially labelled metropolitan areas or metropolitan regions by Eurostat (https://ec.europa.eu/eurostat/web/metropolitan-regions/background, accessed November 25, 2019), for the sake of simplicity we will call them cities in the remaining of the paper, following a large tradition in economics on the relationship between cities, or urban areas, and innovation, starting from Jane Jacobs (Jacobs 1969).

Finally, we exploit the characteristics of the knowledge accessed by looking at qualitative aspects of knowledge generation. In particular, we differentiate between different degrees of knowledge complexity (Sorenson, Rivkin, and Fleming 2006; Balland and Rigby 2017). We expect the advantages of accessing complex knowledge to grow the stronger are the connections between the sender and receiver of the message – that is, when they are socially connected, as opposed to when they only live in the same city. We also account for the quality of the knowledge created through forward citations to patents, a typical indicator for patent quality and value (Jaffe and de Rassenfosse 2017).

Our empirical analysis uses an underexploited database of disambiguated European Patent Office (EPO) inventors residing in Europe, from 1980 to 2010 (Pezzoni, Lissoni, and Tarasconi 2014). Using patent data and the information on inventors, their city of residence, their collaborators, and the firms for which they work listed in patent documents, we build an unbalanced panel at the individual, inventor level. We regress inventors' performance on several measures of the knowledge stock at the level of the city, the firm, and the inventor's network of collaborators. We use the number of patents produced per year as a measure of productivity. Such indicator has limitations but remains the most used measure of inventor productivity (Akcigit et al. 2018; Moretti 2019). Our setting allows us to incorporate a large list of fixed effects (city, firm, time, sector, and individual, plus interactions among them) that rule out the influence of time-invariant confounding factors.

Results highlight the importance of city-level knowledge stocks on inventors' patent production. Doubling the city stocks increases individuals' productivity by 4-5%. We also find positive effects for the firm-level knowledge stocks, but less preponderant in terms of magnitude. Interestingly, when we take into account (above average) patent quality, the picture changes upside down: what matters the most for quality-adjusted patent production is the network and firm-level stocks. In other words, tighter relationships are critical to share the knowledge that allows individuals to produce high-quality ideas.

In the next section, we review the literature outlined above. Section 3 presents our methodological approach, while section 4 describes the data building process and the final dataset. Finally, section 5 outlines the main findings and conclusions follow.

Literature review and hypotheses

Extensive work (e.g., Romer 1990) enabled to appreciate the role of knowledge as the hidden factor boosting firms' productivity thanks to the virtuous presence of externalities or spillovers, i.e. knowledge pieces that can be used by others than the creator at lower than equilibrium-cost. At the very earth of this approach, there are (at least) two pillars: the "special" characteristics of knowledge as a public good – non-rival and only partially appropriable (Arrow 1962), and its recombinant nature (Weitzman 1998). The latter leads to recognizing the interactive and collective nature of knowledge creation, whose generation is, therefore, bounded within the social and geographic limits of interactions between individuals.

The increasing attention upon the bounding and rooted aspects of knowledge dynamics induced a series of empirical studies on Regional Innovation Systems (Asheim and Coenen 2005) and the interplays between various forms of proximity (Boschma 2005). Within these strands of literature, the concept of stock of knowledge developed, as a measure of the knowledge potential of economic agents. Individuals can partly exploit the knowledge of other individuals simply because they are co-located in space, in a bond of institutions, supply chains and repetitive transactions and interactions. The physical constraint gives the chance to build territorial indicators of knowledge that reflect individual endowments by definition, but, in theory, are not the simple sum of individual addends (Antonelli 2000).

Knowledge interactions in cities

Cities are the locus of innovative activity. Cities reduce the costs of interacting and allow accessing and exploiting spatially sticky knowledge (Storper and Venables 2004; Carlino and Kerr 2015; Moretti 2019). As Lucas (1993) puts it, the compact nature of the geographic unit found in cities facilitates communication and interaction, making innovators located in cities more productive than their peers elsewhere. Cities concentrate large quantities of firms and employment which facilitates the spread of knowledge in an unplanned and serendipitous manner, due to geographic proximity (Carlino and Kerr 2015).[3] This is especially so in cities with a large agglomeration of talented people, as

[3] For evidence on the localized nature of knowledge flows, see, among others, Jaffe, Trajtenberg, and Henderson (1993), Peri (2005), Bloom, Schankerman, and Van Reenen (2013) and Singh and Marx (2013).

skilled workers are more able to understand and absorb knowledge spillovers (Lucas 1988).[4] Moreover, cities improve the quality of worker-firm matches (Moretti 2019) and affect labour market efficiency and productivity. Hence, the following hypothesis arises:

H1a. Large, compared to small, cities' knowledge stocks enhance their inventors' productivity.

Network conduits

On a parallel, more recent strand of studies, some authors cast scepticism on the theory of knowledge externalities. The problematic aspect of such a theory is that it treats generation and appropriation of externalities/spillovers as a 'black box', whereas, instead, a multiplicity of forces are at stake (Rodríguez-Pose and Crescenzi 2008). Indeed, knowledge is likely to diffuse through planned and well-structured relations and alliances between individuals, firms and other organizations (Fitjar and Rodríguez-Pose 2017).

Networks are an essential component of innovation because they channel flows of knowledge and information from node to node within a social structure (Owen-Smith and Powell 2004), otherwise more costly to access. Even though geographic co-location and networks might be sometimes observationally equivalent, some authors consider the networks as equally or even more important than the geographic context (Breschi and Lissoni 2001; Breschi and Lissoni 2009).

Onto this track, a literature focusing on networks and regional/firm innovation flourished (Owen-Smith and Powell 2004; Lobo and Strumsky 2008; Miguélez and Moreno 2013; Balland, Belso-Martínez, and Morrison 2016; Breschi and Lenzi 2016; Bergé, Carayol, and Roux 2018; Eriksson and Lengyel 2019). Within this literature, and primarily exploiting the availability of patent data, scholars have uncovered the role of individuals' networks, particularly inventors' networks, as channels of knowledge diffusion (Singh

[4] A debate emerged within this strand of literature on the distinction between specialization externalities (intra-industry) and urbanization externalities (inter-industry). The latter are derived from diversification, which facilitates the exchange of complementary knowledge across different firms and economic agents, yielding to greater returns (Jacobs 1969; Glaeser et al. 1992). Jacobs (1969) stressed that, while Marshallian externalities in clusters/industrial districts are mostly intra-industry, the crucial type of spillovers are across industries, allowing cross-fertilization of ideas.

2005; Breschi and Lissoni 2009) as well as a stimulus for discussions and confrontation of ideas between peers (Bergé, Carayol, and Roux 2018).

To our knowledge, very few studies have analyzed the role of network characteristics on individuals' productivity (Akcigit et al., 2018 is a recent exception). Therefore, we ask whether:

H1b. Large, compared to small, networks' knowledge stocks enhance their inventors' productivity.

The role of the firm

In parallel to these developments, the knowledge-based view of the firm has explored organizations' role in distributing knowledge (Kogut and Zander 1992). This literature contents that the mere existence of the firm facilitates the transferability of knowledge between individuals within firms' boundaries (Allen 1977; Teece 1977; Grant 1996). Tacit knowledge, whose transfer between people is slow, costly and uncertain (Kogut and Zander 1992), particularly benefits from the firm's environment. As knowledge is generally not appropriable through a market transaction, the firm serves as the best platform to organize and share it among the different individuals. That is to say, firms embody the response to the need for coordinating efforts of individual specialists who possess many different types of knowledge and need to share them to produce new ideas (Grant 1996).

Big high-tech firms in Silicon Valley are well aware of the importance of knowledge sharing within the firm and have designed their work spaces to favour interactions among workers. A well-known case is Steve Jobs's design of Pixar Animation Studios, which ensured that engineers, scientists, and executives frequently interacted. Such a perspective has been at the core of many firms' layout designs, including Facebook, Google, Twitter, and AT&T (Carlino and Kerr 2015). Hence, we formulate the following hypothesis:

H1c. Large, compared to small, firms' knowledge stocks enhance their inventors' productivity.

Complex and high-quality ideas

Recent developments in scholarly work are paying increasing attention not only to the quantity of knowledge produced but also to its qualitative aspects. We follow these paths and first differentiate between different degrees of complexity of the knowledge stocks to be accessed (Sorenson, Rivkin, and Fleming 2006; Balland and Rigby 2017). Secondly, we adjust the knowledge produced to its quality (Jaffe and de Rassenfosse 2017).

The cost of acquiring and absorbing knowledge increases with knowledge complexity (Cavusgil, Calantone, and Zhao 2003). Accessing knowledge requires that the receiving partner makes efforts to understand and acquire it, even correcting potential errors in the transmitted message. Thus, complex and highly specific knowledge may diffuse slowly because few agents, apart from the initial innovator, have the necessary capabilities needed to absorb it (Cohen and Levinthal 1990). On the contrary, simple knowledge is accessible even in the case that the sender and the receiver in the social dimension are poorly connected because it does not require the full assistance of the sender to be understood (Sorenson, Rivkin, and Fleming 2006). Barely complex, more routinized forms of knowledge are smoother to move (Balland and Rigby 2017).

We expect the advantages of accessing complex knowledge to grow the stronger are the connections between the sender and receiver of the message – when they are socially or organisationally connected, as opposed to spatially connected only, that is when they live in the same city. The following hypothesis follows:

H2. Knowledge diffusion stemming from knowledge stocks of high complexity is more effective when the sender and the receiver of the flows are at fewer steps of separation – in the same network as opposed to the same firm, and in the same firm as opposed to the same city.

Moreover, inventions differ in their novelty and value (Jaffe and de Rassenfosse 2017; Kogan et al. 2017). Indeed, not all inventions have the same technological and economic impact, and this heterogeneity requires attention. Which type of knowledge interactions does affect the capacity of creating more radical and successful innovations? As illustrated in Kaplan and Vakili (2015), the literature is divided in this issue, and compelling evidence is still lacking. One strand of research states that webs of deeply connected interactions may hurt radical creativity by not allowing researchers to explore new ways of thinking, leading to technological lock-in. Too tight networks, indeed, circulate redundant information since everyone already knows what the others know (Granovetter 1973). Meanwhile, high-quality innovation requires novel recombinations of diverse, distant types of knowledge to be successful (Ahuja and Lampert 2001).

Another stream of research indicates, instead, that, in order to produce higher-than-average ideas, it is necessary to have a profound understanding of the particularities of the knowledge to be accessed (Kaplan and Vakili 2015). Therefore, tight interactions among individuals are more likely to identify anomalies in the knowledge space leading to high-quality innovations (Weisberg 1999; Taylor and Greve 2006; Kaplan and Vakili 2015). Consequently, interactions between inventors within the same firm or in a colleagues' network would allow building new radical knowledge better than those taking place in the city where the inventor lives. Therefore, we put forward the following competing hypotheses:

H3a. High-quality innovations are more likely to appear the broader are knowledge interactions, as wider recombinations are allowed – in the city rather than just the organization, and in the same organization rather than just the network.

H3b. High-quality innovations are more likely to appear the deeper are knowledge interactions, as the capacity to understand and master others' knowledge increases – in the same network as opposed to the same firm, and in the same firm as opposed to the same city.

Empirical strategy

Exploiting patent documents (see section 4), we set our unit of analysis at the inventor level. As introduced by the research hypotheses above, we focus on three social layers, potentially affecting inventors' productivity. They are the network of job relationships each inventor builds around himself during his activity; the firm where s/he is employed; and the city where s/he operates and lives. The three layers cross and sometimes are subsets of each other, but the kind of relationships and exchanges taking place differ, as illustrated above.

We will focus on the stock of knowledge in each layer. In our perspective, the stock of knowledge is not a measure of tangible assets at disposal in knowledge production. Instead, it is an index of an accessible knowledge potential embedded in each layer (the network, the firm or the city). The peculiarity of our approach is that we want to look simultaneously at the different layers building up the inventor's creative environment. There is a long tradition in economics addressing the issue of hierarchical settings, i.e. settings where individuals are nested into groups at many layers (in economics of education see Raudenbush 2009). When the hierarchical structure of the data is ignored, the researcher accepts two implicit assumptions: 1) that the salient heterogeneity takes place only within that layer and that other layers are more or less homogeneous, and 2) that the layer analyzed is independent of the others (Rothaermel and Hess 2007). In some settings, such assumptions may be undesirable or inappropriate. There are two viable approaches to these issues: clustering standard errors in an FE regression setting or running a Multilevel Analysis (MA) (Raudenbush and Bryk 2002; Cameron and Miller 2015). Even though MA has seen only a few applications in regional economics (Fazio and Piacentino 2010; Raspe and van Oort 2011; Tojeiro-Rivero and Moreno 2019), it should generally be preferred over FE (Raudenbush 2009; Bell and Jones 2015; Bell, Fairbrother, and Jones 2016). However, the advantage of the FE estimator is that it eliminates, by definition, group-invariant variables and their interactions with lower-level variables (Clarke et al. 2010). In so doing, any possible correlation between covariates and the errors due to unobserved group-invariant characteristics is avoided. Indeed, one critical point of the MA approach is that the unobserved heterogeneity is not eliminated, meaning that, if the model is not perfectly specified, the omitted variables bias threatens causal interpretation.[5] For all these reasons, we opt for estimating our models by means of FEs.

In order to test our hypotheses, the following FEs regression is going to be estimated:

$$\begin{aligned} InvProd_{\,i,f,c,t} = {} & \beta_1 \cdot City\ Stock_{\,c,t-1} + \beta_2 \cdot Firm\ Stock_{\,f,t-1} \\ & + \beta_3 \cdot Network\ Stock_{\,i,t-1} + \beta_n \cdot Controls_{\,i,f,c,t-1} + \delta_{\,i} + \delta_{\,f} \quad (1) \\ & + \delta_{\,c} + \delta_{\,t} + \delta_{\,gt} + \delta_{\,gc} + \varepsilon_{\,i,f,c,t} \end{aligned}$$

We regress inventors' productivity on the stock of knowledge of the city, firm and network of past collaborators. *i* is the inventor, *f* the firm, *c* the city, *t* stands for time, *g* is

[5] Both Raudenbush (2009) and Bell and Jones (2015) suggest a robust version of MA, where variables are demeaned as in the Mundlack formulation of the FE estimator. When more than one group fixed effect is needed, sequential demeaning is allowed in balanced panels However, the validity of this approach in unbalanced panels has not been tested.

the technology, and the δs are a set of FEs. Inventor, firm, and city effects account for sorting due to permanent differences in productivity across workers, firms and cities. Technology-per-time FEs, δ_{gt}, are present in order to account for technology-specific shocks that may drive inventors' patenting across cities and firms. We also add technology-city FEs, δ_{gc}, in order to absorb time-invariant confounders specific to a technology-city pair (e.g., proximity to a university skilled in a specific technology, like the CERN in Geneva).

The three main explanatory variables account for the knowledge potential of the multilevel structure the inventor is embedded in. In our specification, we address the multilevel structure of the data with a full set of interactions between the main variables of interest across levels and cluster-robust SE (Cameron and Miller 2015).

Data

We match two different patent databases in order to retrieve all the necessary information about our three layers of interest: the PATENTS-ICRIOS database (Coffano and Tarasconi 2014) and the OECD HAN Database (2016). Both databases have the crucial feature of being the output of a process of name-disambiguation: inventors' names in ICRIOS, through inventors' IDs assigned by Pezzoni, Lissoni, and Tarasconi (2014), and applicants' names in HAN. Out of this matched dataset, we build our main variables of interest: the number of patent applications per inventor-year (dependent variable), and a series of knowledge (patent) stocks for i) the inventor's network of collaborators in a five-year window, ii) the firm-year and iii) the city-year tuples.

Each patent is assigned to a layer (the network, the firm or the city) with a whole count. It means that if a patent application is assigned to more than one applicant, the stock count of each of these applicants increases in one unit rather than a proportion of it – as it would be for fractional counts. One peculiar characteristic of knowledge, i.e. knowledge indivisibility, supports this approach. After the assignation to each repository, the patent stock is discounted every year with a 15% depreciation factor (the so-called Permanent Inventory Method, see Hall, Jaffe, and Trajtenberg, 2005)

Dependent variables

Inventors' patent production. The dependent variable of our baseline model is a bare count of the yearly patent applications signed by an inventor. We only observe non-zero counts; hence we exploit the variance in the size of each inventor's production. Moreover, in order to control for individual time-invariant unobserved characteristics, we retain only inventors who invented at least twice.

Inventors' quality-adjusted patent production. As an alternative dependent variable, we also computed the count of high-quality yearly patent applications signed by an inventor. High-quality patents are the top-50% patents sorted by their forward citations received – within a time window of 5 years after the priority year of the cited patent – controlling for technological area and cohort (Waltman et al. 2011; Wohlrabe and Bornmann 2017). Citations data comes from PATENTS-ICRIOS database. We use DOCDB families to count forward citations but avoid double-counting.

Explanatory variables

City. We identify our city boundaries using EUROSTAT's definition of "Metropolitan Regions", which correspond to "NUTS 3 regions or a combination of NUTS 3 regions which represent all agglomerations of at least 250,000 inhabitants. These agglomerations were identified using the Urban Audit's Functional Urban Area (FUA)".[6] FUA identifies a city of >250,000 inhabitants plus its commuting zone. We add some areas excluded by the original definition but relevant according to patent production rates. Indeed, we retained FUAs whose yearly patent production is equal to that of cities belonging to the upper quartile of cities' patent distribution (e.g. Cambridge Area). In order to locate inventors in these cities and compute cities' knowledge stocks, we match our databases to the OECD REGPAT Database (Maraut et al. 2008) which provides regionalized information (NUTS3 level for Europe) for all EPO inventors.

Firm. We use the applicant's name – the owner of the patent – listed in patent documents as a proxy for the company (or other organizations such as universities or research centres) the inventor works for. Firm names are from the OECD HAN Database (2016), which exploits ORBIS for applicants' name harmonization. Some applicants are multi-

[6] See https://ec.europa.eu/eurostat/web/metropolitan-regions/background for full information (accessed September 22, 2020).

establishment companies, and some of these establishments are located in different cities. Henceforth, we treat same-city establishments as one and define it as a firm. Consequently, establishments of the same company but located in different cities will be different firms. Given that in most patent documents the headquarter's address is reported only, we identify establishments thanks to the geo-localization of inventors' addresses employed by that company (applicant).

Network. To define the collaboration network, we consider those inventors who worked with the focal inventor within a 5-year window up to one year before the focal year. We assume that a past collaboration remains an active source of knowledge for a 5-year period at most (Breschi and Lenzi 2016). Finally, we sum up the individuals' depreciated stocks of the collaborators.

Complexity

We qualify each layer's knowledge stock with a measure of modular complexity, as introduced in Fleming and Sorenson (2001, 2004). This measure relies on the conceptualization of invention as a search in a technological knowledge landscape. Landscapes are made of components, which in turn are measured by technological classes listed in the patent document. The position in the landscape represents a combination of components, with an associated fitness value. Creativity takes the form of a movement on the landscape until a position with a higher fitness appears. The outcome of the search process depends on one factor: the interdependence among technological components. The concept of interdependence coincides with that of modularity or coupling, that is, when two entities are interdependent, a small change in one component calls for changes in the other component in order to the combination to work correctly. We operationalize such procedure as follow.

a. The *Ease of Recombination* (*EoR*) is computed for each technological class-year of the dataset, being technological classes specified as 4-digit IPCs. The *EoR* is the ratio between the count of classes previously combined with the focal class, and the number of applications referencing to the focal class.
b. We calculate *Modular Complexity* (*MC*) index for each patent application as the count of technological classes of the focal patent divided by the sum of their *EoR*, i.e. the inverse of a weighted average.

After computing such index for every patent, we assign to "low" complexity the applications belonging to the lower 50% of the distribution, to "medium" those belonging to the upper 50-to-90% of the distribution, and "high" the remaining ones. Finally, we compute three separate stocks for each level of modular complexity at each layer. We also apply a normalization procedure following Alstott et al. (2017) (see the online Appendix 1 for more details).

All variables enter the regression models after an Inverse Hyperbolic Sine (IHS) transformation, which is a log-like transformation well-defined at zero (differently from the natural logarithm).[7] Moreover, knowledge stocks come with a one year lag to allow the search and absorption of knowledge by inventors to take place.

Controls

We want that the variables measuring the stock of knowledge at the different layers grasp knowledge capacity and accessibility only, rather than the intensity of innovativeness. Therefore, we compute a set of layer-based control variables for productivity, measured as the average inventor's productivity at each of these layers.

In order to keep at a minimum the correlation with the stock variables, instead of the bare count, we use the quality-adjusted count (see above). We proceed as follows: first, we compute individual inventor's productivity on a time window between *t* and *t-4*. Consequently, averages are computed at each layer. The network's productivity is computed on a 5-year window from *t-1* backwards, whereas the firm-level variable is computed on a 3-year window from *t-1* backwards in order to minimize missing values in the lag variable (very few firms invent more than once in consecutive years). City-level average productivity enters the regressions as the respective value for each city at *t-1*.

Evidence of the importance of multinationals (MNEs) in affecting firms, territorial productivity and knowledge capacity is growing (Iammarino and McCann 2013). As stated by Crescenzi, Gagliardi, and Iammarino (2015), MNEs are amongst the leading creators of new technology worldwide, as well as its transfer and diffusion in the world economy. Therefore, we control for the multinational effect with a dummy coded one if

[7] See MacKinnon and Magee (1990) and the post at https://worthwhile.typepad.com/worthwhile_canadian_initi/2011/07/a-rant-on-inverse-hyperbolic-sine-transformations.html for more information (accessed September 22, 2020).

the applicant is present in more than one country in our dataset – both if they are inventing in more than one nation, and if their employed inventors declare to live in countries different from that of the firm's headquarter.

Even though we aim at controlling for other relevant socio-economic regional variables, most of them are not available for a long time window at our territorial unit of analysis. The Cambridge Econometrics (CE) Database partially provides NUTS3 level data on GVA, population and employment, which we use to compute gross value-added per capita (GVA pc), population density and a Herfindahl-Hirschman index of employment specialization. We show regressions with CE controls in the online Appendix 5.

The inventor-firm-city matching process generated multiple ambiguous assignations, e.g. more than one city or firm for inventor-year. We set up some decision rules for disambiguating such matches. The rationale is continuity, i.e. we want to detect when mobility patterns of inventors across firms and cities are too frequent to be realistic, and we assign more weight to long-lasting ties in case of plausible ambiguous assignations (Hoisl 2007; Nakajima, Tamura, and Hanaki 2010).[8]

Our final dataset results in a strongly unbalanced panel of ~818.000 observations, clustered in

- ~243.000 multiple inventors (inventors that applied for patents more than once)
- ~60.000 applicants (firms);
- 318 cities;
- 31 years, from 1980 to 2010.

Descriptive evidence

Figure 1 shows the cities considered in the present study, coloured according to their level of knowledge stock computed in 1980 and 2010 (the two extremes of our timeframe). As can be seen, the cities with the most extensive knowledge stocks are in the core of Europe. Some cities (particularly in the South of Europe) seem to converge in terms of knowledge stocks (they scale up to belong to the group with the largest stocks) but most cities show a constant figure over time.

[8] More details are presented in the online Appendix 1.

[Insert Figure 1 about here]

Table 1 lists the top-10 firms considered in our study, sorted by their knowledge stocks in 2010. Top firms are typically large multinational companies, in some cases showing up several times in the top list as a consequence of their multi-city presence.

[Insert Table 1 about here]

Table A.1 in the online Appendix provides the descriptive statistics of our sample, while Table A.2 presents the correlation matrix. Correlation is low for all pairs except for the network knowledge stock and the network productivity variables (0.78). However, simple correlations might not be adequate in a panel data framework to gauge multicollinearity problems, and therefore we run Variance Inflation Factor (VIF) tests after regressions. Fortunately, these do not point to severe collinearity problems (results provided upon request from the authors).

Results

Table 2 reports the results of the FEs regressions. From column 1, we learn that the effect of city-level knowledge stocks is positive and significant. Doubling the size of the stock augments productivity by 4.8% – results from IHS transformed variables can be interpreted as elasticities. These results are not far from the ones found in economic geography when estimating agglomeration effects (Rice, Venables, and Patacchini 2006).

The following columns introduce explanatory variables in a cascade. Column 2 introduces firm-level stocks, which are positive and significant, but to a lesser extent than the city-level ones. Network-level stock enters in column 3. Simultaneous consideration of the three relevant knowledge flows levels let us better gauge the respective coefficients and significance than previous researches. In terms of magnitude, city-level knowledge stocks are still much more potent than network-level ones: city-level stocks' regression coefficient shrinks but still keeps its lead. Column 4 introduces the remaining relevant controls, and some interesting findings emerge. First, the role of city-level stocks remain positive and significant, and with similar size as compared to estimates without controls. Inventors also take advantage of the firm-level knowledge pools, although the elasticity is relatively small compared to the city-level one. Finally, once we account for productivity controls, the network-level stock becomes not statistically significant. The negative and significant sign of the city productivity marks that city peers' innovativeness tends to inhibit inventors capacity to produce new knowledge, i.e. a downside of competition. All in all, hypotheses H1a and H1b are confirmed, while H1c is rejected.

[Insert Table 2 here]

We next assess the validity of the assumption of independence across levels by introducing interaction effects between city-, firm- and network-level stocks. In particular, we pursue the question of whether the different levels taken into account are complementary or substitutive to each other, by plotting marginal effects of one of the stock variables at quartiles of the interacted one. We find that firm and network stocks reinforce one another at the margin, indicating positive feedbacks between creative and generative work environments and resourceful collaboration networks (Figure 2.A). In other words, the inventor's ability to exploit efficiently the external knowledge coming from his/her networks also depends upon the knowledge capacity of his/her work environments.[9] It is important to note that past collaborators may overlap with current workplace colleagues, but only partially: the inventor's network may stretch well beyond firm and city boundaries.

On the contrary, network and city stocks are substitutes (Figure 2.B): we observe diminishing returns at the margins in the use of the network stock when the city stock is high and vice-versa. Finally, the city-firm interaction term is not significant; therefore,

[9] Lane and Maxfield (2005) introduce the concept of generativity in a relationship facing radical uncertainty as knowledge production. The generativity of the relationship, they say, stems from the empowered capacity of the interacting individuals to build new interpretative structure of reality, i.e. to see old things with new eyes. Generativity has five prerequisites: some commonality of intentions, benevolence in approaching the other's differences, confidence to share doubts and inspirations, and some heterogeneity – whatever kind it is, not necessarily in terms of competence or skills.

we do not plot it. However, regression results are shown in the online Appendix (Table A.3).

One possible reading key for these interaction effects concerns the characteristics of the knowledge prevalently exchanged at each layer. We might expect that knowledge flowing at the city-level is generic and accessible so that it can pass through sparse and fragmented interactions, whereas the one exchanged through the network is specific. As for the firm, it is a mixture of the two, but what matters is that it is a generative, resourceful environment where specific knowledge can be metabolized. By adding generic knowledge onto specific knowledge, it comes redundancy. Nevertheless, by providing specific knowledge in a generative environment, the result is higher efficiency. Finally, there is no evidence that more generic knowledge in a generative environment does add anything.

[Insert Figure 2 about here]

Figure 3 plots coefficients from regressions splitting the knowledge stocks in each layer according to their degree of complexity (low, medium and high). Black bars correspond to column 4 of Table 2 for comparison purposes. Results indicate that for low and medium levels of complexity, results are similar to previous tables: effects for city-level stocks are larger than firm-level stocks, whereas network-level stocks effects are not statistically significant. However, some differences are worth reporting. The network-level coefficient of low complexity is actually statistically significant but negative. It means that when inventors access low complexity knowledge through their network, they are wasting resource (e.g. time) in the wrong transmission channel. It may happen because they can collect the same kind of knowledge through the city and firm-levels and if they do, resorting to the network provides redundant knowledge and incurs in a congestion effect. On the contrary, for high levels of complexity, knowledge resists diffusion when inventors access it at the city- and network-levels (not statistically significant coefficients). At the same time, the coefficient remains strongly significant at the firm-level, evidencing the importance of close interactions within organizational boundaries to transfer complex ideas, where possibly onsite demonstrations, direct monitoring and learning-by-doing is the rule. Admittedly, we would have expected the network to be a proficient vehicle of high complexity knowledge because it allows specific and even tacit knowledge to circulate. Therefore, hypothesis H2 is only partially confirmed.

[Insert Figure 3 about here]

Our expectations are confirmed when we look at quality-adjusted knowledge production. As anticipated in the introductory section, results reported in Figure 4 change upside down the story above. Considering the black bars first (baseline without splitting by complexity levels), we learn that city-level knowledge stocks do not have any effect on the number of high-quality patents the inventors produce. The result suggests that if inventors aim to produce high-quality ideas, they must rely on direct and knitted ties. Indeed, the network-level is taking the lead now (the firm-level stock's coefficient is only mildly significant at 95% only). Thus, a search over the knowledge space guided by circumscribed and directed professional relationships is more likely to identify anomalies leading to high-quality innovations. The message to be transmitted in such cases is usually more tacit, and therefore closer interactions and trust between the sender and the receiver of the messages are crucial to generate high-quality ideas. Thus, hypothesis H3a is rejected, while H3b is confirmed.

[Insert Figure 4 about here]

By splitting the knowledge stocks by their degree of complexity (low, medium and high), the results appreciate the role of the firm-level stocks too, but the network-level still predominates in terms of size for high complexity knowledge transmission.

In Appendix 5, Table A.6 reports some robustness checks. We substituted the local firm-level variables with global firm-level ones (column 1). Results are very similar to those of Table 2. In column 2, we add CE controls for GVA per capita (economic performance), population density (agglomeration economies) and employment HH index (industrial specialization), but none of them turns statistically significant, and results keep unchanged.

Conclusions

Understanding the mechanisms of knowledge creation and diffusion is a relevant topic in economics and other social sciences since knowledge creation is at the very base of the innovative dynamics and behind the economic growth of firms, cities and countries. The present research contributes to enriching such understanding, appreciating the complexity of the social structure where innovators are embedded in. We provide a novel contribution in many respects. First, we analyze individual inventors' capacity to create new knowledge with a large, longitudinal dataset. In so doing, we apply the precious inheritance of the economics of knowledge as well as economic geography, mainly dealing with regions and firms, to individuals. Second, we exploit the new EUROSTAT classification of the European territories by Metropolitan Regions in order to target more efficiently than before the actual geographical locus of knowledge production. Third, we account simultaneously for what we believe are the three most fundamental levels where knowledge flows: the city, the firm, and the individual's network.

Combing those contributions together, we can uncover several compelling results. First and foremost, even after accounting for the knowledge potential delivered by the network of collaborators, and the knowledge capacity of the firm where inventors do invent, city-level knowledge diffusion stands up as a significant and sizeable force enhancing knowledge production – the most effective layer. Doubling the city stock increases individuals' productivity by 4-5%. Although not directly comparable, it is worth highlighting that these estimates are in line with the effects of agglomeration economies on wages and in line with the vast literature on the territorial aspects of innovation we partially reviewed above. It may be that city knowledge, albeit being mostly generic and accessible, induce the largest productivity premium because of its variety.

Second, the firm-level knowledge capacity emerges as pivotal: it not only exerts a positive standalone effect on inventors' productivity, but it also proves crucial for the effectiveness of the network knowledge. Firms, indeed, are a middle layer, branding with some of both individual and geographic-level characteristics, but, besides, they qualify with learning-by-doing. It is, probably, in this kind of learning scheme that the specific, tacit knowledge sourced from the collaboration network blossoms and is magnified.

When ideas' quality is considered, as well as the type of knowledge accessed to produce them, the picture changes considerably. While city-level knowledge stocks of low and medium complexity influence inventors' productivity, high-complexity knowledge stocks are productive when coming from the inventors' firm only. The message is, complex ideas are productive only if inventors are bound enough to be able to understand and assimilate such kind of messages and transform them into patentable artefacts. Finally, when it comes to producing above-average quality ideas, the network stocks stand out, suggesting that closer, person-to-person relations are critical. The effect is even reinforced when knowledge stocks qualify by their degree of complexity: the effect of network stock on high-quality innovations are especially large when this knowledge is of medium and high complexity. Tacit knowledge flowing by virtue of the confidentiality and trust built within the network, together with the interactive learning scheme emerging through direct collaborations, help to explain the whys for these findings.

We are aware of course that our regressions do not entirely tackle endogeneity issues, mostly related to the omission of relevant variables. We partially attenuated such omission with an extensive list of FE, clearing regression results from time-invariant confounding effects. Still, time-varying confoundings may exert a relevant effect; however, data sources covering the three layers for such a long time and vast geographic space are not available. Our future research aims at addressing this shortcoming, although it may require to focus on a single layer at a time. In this research, we focused on the multilevel approach, not well-developed in the literature, and we exploited patent data at its maximum. The vast range of empirical exercises we provided indicates that the multilevel approach is fertile and probably the appropriate path to follow for future assessments of knowledge dynamics.

Data availability statement

The data that support the findings of this study are openly available https://figshare.com/s/64d8dd730eacaec000ea.

References

Ahuja, G., and C.M. Lampert. 2001. Entrepreneurship in the Large Corporation: A Longitudinal Study of How Established Firms Create Breakthrough Inventions. *Strategic Management Journal* 22, no. 6-7: 521–543.

Akcigit, U., S. Caicedo, E. Miguelez, S. Stantcheva, and V. Sterzi. 2018. *Dancing with the Stars: Innovation through Interactions*. National Bureau of Economic Research.

Akçomak, İ.S., and B. ter Weel. 2009. Social Capital, Innovation and Growth: Evidence from Europe. *European Economic Review* 53, no. 5 (July): 544–567.

Allen, T.J. 1977. Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information within the R & D Organization(Book). *Research Supported by the National Science Foundation. Cambridge, Mass., MIT Press, 1977. 329 p*.

Alstott, J., G. Triulzi, B. Yan, and J. Luo. 2017. Mapping Technology Space by Normalizing Patent Networks. *Scientometrics* 110, no. 1: 443–479.

Andersson, M., and C. Karlsson. 2002. The Role of Accessibility for Regional Innovation Systems.

———. 2004. 10. The Role of Accessibility for the Performance of Regional Innovation Systems. *Knowledge Spillovers and Knowledge Management* 283.

Antonelli, C. 2000. Collective Knowledge Communication and Innovation: The Evidence of Technological Districts. *Regional Studies* 34, no. 6 (August): 535–547.

Arrow, K. 1962. Economic Welfare and the Allocation of Resources for Invention. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, 609–626. Princeton University Press. http://www.nber.org/chapters/c2144.pdf.

Asheim, B.T., and L. Coenen. 2005. Knowledge Bases and Regional Innovation Systems: Comparing Nordic Clusters. *Research Policy* 34, no. 8 (October): 1173–1190.

Balland, P.-A., J.A. Belso-Martínez, and A. Morrison. 2016. The Dynamics of Technical and Business Knowledge Networks in Industrial Clusters: Embeddedness, Status, or Proximity? *Economic Geography* 92, no. 1 (January 2): 35–60.

Balland, P.-A., and D. Rigby. 2017. The Geography of Complex Knowledge. *Economic Geography* 93, no. 1 (January): 1–23.

Bell, A., M. Fairbrother, and K. Jones. 2016. Fixed and Random Effects: Making an Informed Choice. *Under Review*. https://www.researchgate.net/profile/Andrew_Bell7/publication/299604336_Fixed_and_Random_effects_making_an_informed_choice/links/57028c1608aeade57a246326/Fixed-and-Random-effects-making-an-informed-choice.pdf.

Bell, A., and K. Jones. 2015. Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. *Political Science Research and Methods* 3, no. 01 (January): 133–153.

Bergé, L., N. Carayol, and P. Roux. 2018. How Do Inventor Networks Affect Urban Invention? *Regional Science and Urban Economics* 71: 137–162.

Bloom, N., M. Schankerman, and J. Van Reenen. 2013. Identifying Technology Spillovers and Product Market Rivalry. *Econometrica* 81, no. 4: 1347–1393.

Boschma, R. 1999. Learning and Regional Development. *GeoJournal* 49, no. 4: 339–343.

———. 2005. Proximity and Innovation: A Critical Assessment. *Regional Studies* 39, no. 1 (February): 61–74.

Boschma, R., P.-A. Balland, and D.F. Kogler. 2015. Relatedness and Technological Change in Cities: The Rise and Fall of Technological Knowledge in US Metropolitan Areas from 1981 to 2010. *Industrial and Corporate Change* 24, no. 1 (February 1): 223–250.

Boschma, R., and K. Frenken. 2010. The Spatial Evolution of Innovation Networks. A Proximity Perspective. *Handbook of Evolutionary Economic Geography*: 120–135.

Breschi, S., and C. Lenzi. 2016. Co-Invention Networks and Inventive Productivity in US Cities. *Journal of Urban Economics*.

Breschi, S., and F. Lissoni. 2001. Localised Knowledge Spillovers vs. Innovative Milieux: Knowledge “Tacitness” Reconsidered. *Papers in Regional Science* 80, no. 3 (July 1): 255–273.

———. 2009. Mobility of Skilled Workers and Co-Invention Networks: An Anatomy of Localized Knowledge Flows. *Journal of Economic Geography* 9, no. 4: 439–468.

Cameron, A.C., and D.L. Miller. 2015. A Practitioner’s Guide to Cluster-Robust Inference. *Journal of Human Resources* 50, no. 2: 317–372.

Carlino, G., and W.R. Kerr. 2014. *Agglomeration and Innovation*. Working Paper. National Bureau of Economic Research. http://www.nber.org/papers/w20367.

Carlino, G., and W.R. Kerr. 2015. Agglomeration and Innovation. In *Handbook of Regional and Urban Economics*, 5:349–404. Elsevier.
Cavusgil, S.T., R.J. Calantone, and Y. Zhao. 2003. Tacit Knowledge Transfer and Firm Innovation Capability. *The Journal of Business and Industrial Marketing* 18, no. 1: 6–21.
Clarke, P., C. Crawford, F. Steele, and A.F. Vignoles. 2010. The Choice between Fixed and Random Effects Models: Some Considerations for Educational Research. *The Centre for Market and Public Organisation 10/240*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1700456.
Coffano, M., and G. Tarasconi. 2014. CRIOS-Patstat Database: Sources, Contents and Access Rules. *Center for Research on Innovation, Organization and Strategy, CRIOS Working Paper* 1, no. 1.
Cohen, W.M., and D.A. Levinthal. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly* 35, no. 1 (March): 128.
Crescenzi, R., L. Gagliardi, and S. Iammarino. 2015. Foreign Multinationals and Domestic Innovation: Intra-Industry Effects and Firm Heterogeneity. *Research Policy* 44, no. 3: 596–609.
Duranton, G., and D. Puga. 2001. Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. *American Economic Review* 91, no. 5 (December): 1454–1477.
Eriksson, R.H., and B. Lengyel. 2019. Co-Worker Networks and Agglomeration Externalities. *Economic Geography* 95, no. 1 (January 1): 65–89.
Fazio, G., and D. Piacentino. 2010. A Spatial Multilevel Analysis of Italian SMEs' Productivity. *Spatial Economic Analysis* 5, no. 3 (September): 299–316.
Fitjar, R.D., and A. Rodríguez-Pose. 2017. Nothing Is in the Air. *Growth and Change* 48, no. 1: 22–39.
Fleming, L., and O. Sorenson. 2004. Science as a Map in Technological Search. *Strategic Management Journal* 25: 909–928.
Fleming, Lee. 2001. Recombinant Uncertainty in Technological Search. *Manage. Sci.* 47, no. 1 (January): 117–132.
Fleming, Lee, and O. Sorenson. 2001. Technology as a Complex Adaptive System: Evidence from Patent Data. *Research Policy* 30, no. 7: 1019–1039.
Giuri, P., and M. Mariani. 2013. When Distance Disappears: Inventors, Education, and the Locus of Knowledge Spillovers. *Review of Economics and Statistics* 95, no. 2: 449–463.
Glaeser, E.L., H.D. Kallal, J.A. Scheinkman, and A. Shleifer. 1992. Growth in Cities. *Journal of Political Economy* 100, no. 6: 1126–52.
Granovetter, M.S. 1973. The Strength of Weak Ties. *American Journal of Sociology*: 1360–1380.
Grant, R.M. 1996. Toward a Knowledge-Based Theory of the Firm: Knowledge-Based Theory of the Firm. *Strategic Management Journal* 17, no. S2 (December): 109–122.
Hall, B.H., A. Jaffe, and M. Trajtenberg. 2005. Market Value and Patent Citations. *RAND Journal of Economics* 36, no. 1: 16–38.
Hoisl, K. 2007. Tracing Mobile Inventors—The Causality between Inventor Mobility and Inventor Productivity. *Research Policy* 36, no. 5 (June): 619–636.
Iammarino, S., and P. McCann. 2013. *Multinationals and Economic Geography: Location, Technology and Innovation*. Cheltenham: Edward Elgar.
Jacobs, J. 1969. Economy of Cities.
Jaffe, A.B., and G. de Rassenfosse. 2017. Patent Citation Data in Social Science Research: Overview and Best Practices. *Journal of the Association for Information Science and Technology* 68, no. 6 (June 1): 1360–1374.
Jaffe, A.B., M. Trajtenberg, and R. Henderson. 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics* 108, no. 3 (August 1): 577–598.
Jaravel, X., N. Petkova, and A. Bell. 2018. Team-Specific Capital and Innovation. *American Economic Review* 108, no. 4–5 (April): 1034–1073.

Kaplan, S., and K. Vakili. 2015. The Double-edged Sword of Recombination in Breakthrough Innovation. *Strategic Management Journal* 36, no. 10 (October 1): 1435–1457.
Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman. 2017. Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics* 132, no. 2 (May 1): 665–712.
Kogut, B., and U. Zander. 1992. Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science* 3, no. 3 (August): 383–397.
Lambooy, J.G. 1997. Knowledge Production, Organisation and Agglomeration Economies. *GeoJournal* 41, no. 4: 293–300.
Lane, D.A., and R.R. Maxfield. 2005. Ontological Uncertainty and Innovation. *Journal of Evolutionary Economics* 15, no. 1: 3–50.
Lobo, J., and D. Strumsky. 2008. Metropolitan Patenting, Inventor Agglomeration and Social Networks: A Tale of Two Effects. *Journal of Urban Economics* 63, no. 3 (May): 871–884.
Lucas, R.E. 1988. On the Mechanics of Economic Development. *Journal of Monetary Economics* 22, no. 1 (July): 3–42.
Lucas, R.E. 1993. Making a Miracle. *Econometrica: Journal of the Econometric Society*: 251–272.
Lucas, R.E. 2009. Ideas and Growth. *Economica* 76, no. 301 (February 1): 1–19.
Lucas, R.E., and B. Moll. 2014. Knowledge Growth and the Allocation of Time. *Journal of Political Economy* 122, no. 1: 1–51.
Lundvall, B.-Å. 1988. Innovation as an Interactive Process: From User Producer Interaction to National Systems of Innovation. In *Technical Change and Economic Theory*.
Lundvall, B.-A. 1996. The Social Dimension of the Learning Economy.
MacKinnon, J.G., and L. Magee. 1990. Transforming the Dependent Variable in Regression Models. *International Economic Review*: 315–339.
Marshall, A. 1890. Principles of Economics, (New York, 1948). *Marshall8229Principles of Economics1948*: 229.
Maskell, P., and A. Malmberg. 1999. The Competitiveness of Firms and Regions: 'Ubiquitification'and the Importance of Localized Learning. *European Urban and Regional Studies* 6, no. 1: 9–25.
Miguélez, E., and R. Moreno. 2013. Research Networks and Inventors' Mobility as Drivers of Innovation: Evidence from Europe. *Regional Studies* 47, no. 10: 1668–1685.
Miguélez, E., R. Moreno, and M. Artís. 2011. Does Social Capital Reinforce Technological Inputs in the Creation of Knowledge? Evidence from the Spanish Regions. *Regional Studies* 45, no. 8: 1019–1038.
Moretti, E. 2019. *The Effect of High-Tech Clusters on the Productivity of Top Inventors*. Working Paper. National Bureau of Economic Research. http://www.nber.org/papers/w26270.
Nakajima, R., R. Tamura, and N. Hanaki. 2010. The Effect of Collaboration Network on Inventors' Job Match, Productivity and Tenure. *Labour Economics* 17, no. 4 (August): 723–734.
Nooteboom, B. 2000. Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance. *Journal of Management and Governance* 4, no. 1 (March 1): 69–92.
Owen-Smith, J., and W.W. Powell. 2004. Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community. *Organization Science* 15, no. 1 (January 1): 5–21.
Peri, G. 2005. Determinants of Knowledge Flows and Their Effect on Innovation. *Review of Economics and Statistics* 87, no. 2 (May 1): 308–322.
Pezzoni, M., F. Lissoni, and G. Tarasconi. 2014. How to Kill Inventors: Testing the Massacrator© Algorithm for Inventor Disambiguation. *Scientometrics* 101, no. 1: 477–504.
Raspe, O., and F.G. van Oort. 2011. Firm Heterogeneity, Productivity and Spatially Bounded Knowledge Externalities. *Socio-Economic Planning Sciences* 45, no. 1 (March): 38–47.

Raudenbush, S.W. 2009. Adaptive Centering with Random Effects: An Alternative to the Fixed Effects Model for Studying Time-Varying Treatments in School Settings. *Education Finance and Policy* 4, no. 4: 468–491.
Raudenbush, S.W., and A.S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. Advanced Quantitative Techniques in the Social Sciences 1. Thousand Oaks: Sage Publications.
Rice, P., A.J. Venables, and E. Patacchini. 2006. Spatial Determinants of Productivity: Analysis for the Regions of Great Britain. *Regional Science and Urban Economics* 36, no. 6: 727–752.
Rodríguez-Pose, A., and R. Crescenzi. 2008. Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe. *Regional Studies* 42, no. 1 (February): 51–67.
Romer, P.M. 1990. Endogenous Technological Change. *Journal of Political Economy* 98, no. 5 (October 1): S71–S102.
Rothaermel, F.T., and A.M. Hess. 2007. Building Dynamic Capabilities: Innovation Driven by Individual-, Firm-, and Network-Level Effects. *Organization Science* 18, no. 6 (December): 898–921.
Singh, J. 2005. Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science* 51, no. 5 (January 5): 756–770.
Singh, J., and M. Marx. 2013. Geographic Constraints on Knowledge Spillovers: Political Borders vs. Spatial Proximity. *Management Science* 59, no. 9 (April 4): 2056–2078.
Sorenson, O., J.W. Rivkin, and L. Fleming. 2006. Complexity, Networks and Knowledge Flow. *Research Policy* 35, no. 7 (September): 994–1017.
Storper, M., and A.J. Venables. 2004. Buzz: Face-to-Face Contact and the Urban Economy. *Journal of Economic Geography* 4, no. 4 (August 1): 351–370.
Taylor, A., and H.R. Greve. 2006. Superman or the Fantastic Four? Knowledge Combination and Experience in Innovative Teams. *Academy of Management Journal* 49, no. 4: 723–740.
Teece, D.J. 1977. Technology Transfer by Multinational Firms: The Resource Cost of Transferring Technological Know-How. *The Economic Journal* 87, no. 346: 242–261.
Tojeiro-Rivero, D., and R. Moreno. 2019. Technological Cooperation, R&D Outsourcing, and Innovation Performance at the Firm Level: The Role of the Regional Context. *Research Policy* 48, no. 7: 1798–1808.
Waltman, L., N.J. van Eck, T.N. van Leeuwen, M.S. Visser, and A.F.J. van Raan. 2011. Towards a New Crown Indicator: Some Theoretical Considerations. *Journal of Informetrics* 5, no. 1 (January): 37–47.
Weisberg, R.W. 1999. I2 Creativity and Knowledge: A Challenge to Theories. *Handbook of Creativity* 226.
Weitzman, M.L. 1998. Recombinant Growth. *The Quarterly Journal of Economics* 113, no. 2: 331–360.
Wohlrabe, K., and L. Bornmann. 2017. Normalization of Citation Impact in Economics. *MPRA Paper No. 80384*. https://mpra.ub.uni-muenchen.de/80384/.

Figures and tables

Figure 1. Cities/metro areas considered and their distribution of knowledge stocks

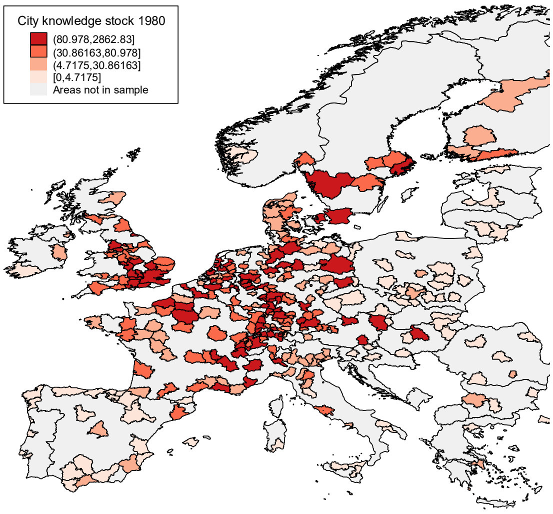


Figure 1.A: City-level knowledge stocks 1980

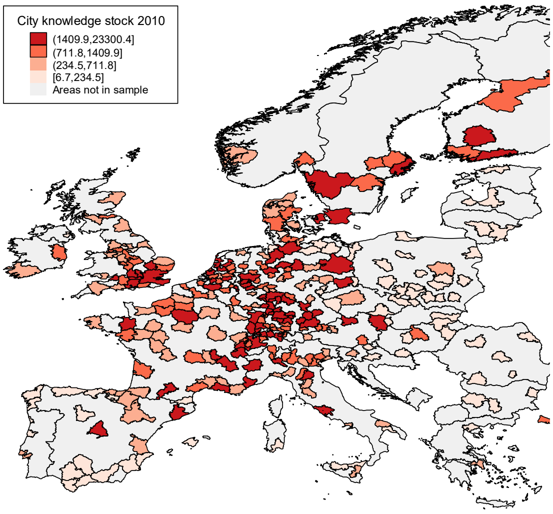


Figure 1.B: City-level knowledge stocks 2010

Figure 2. Predicted margins of interactions between main variables of interest

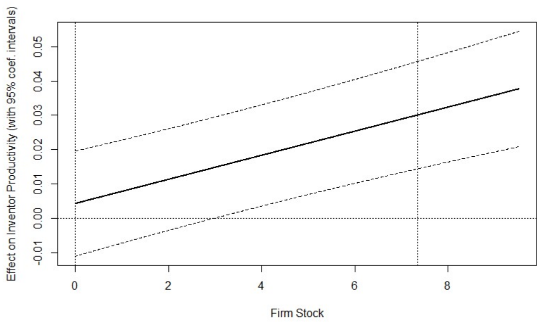


Figure 2.A: Effect of network stock on inventor productivity per levels of firm stocks

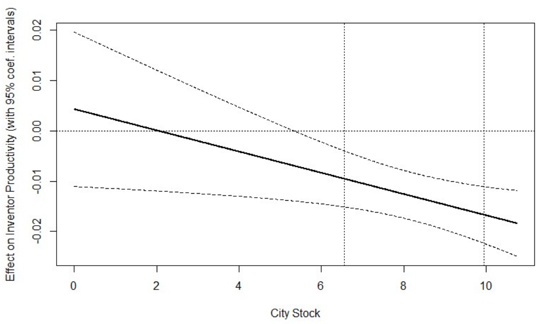
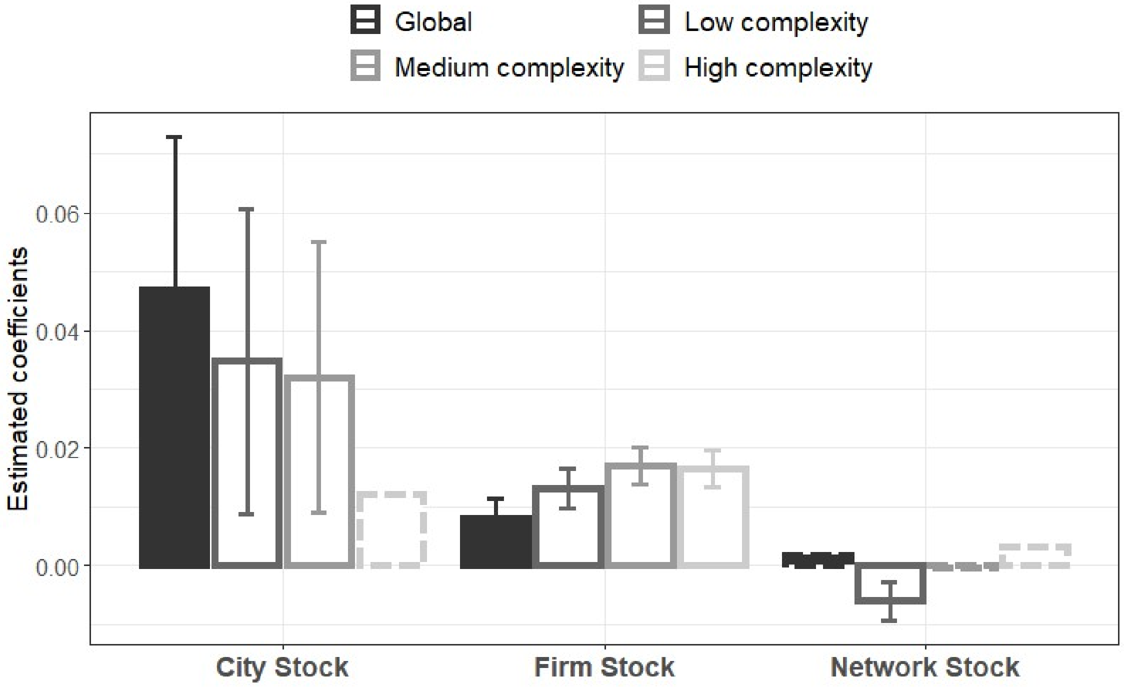


Figure 2.B: Effect of network stock on Inventor productivity per levels of city stocks

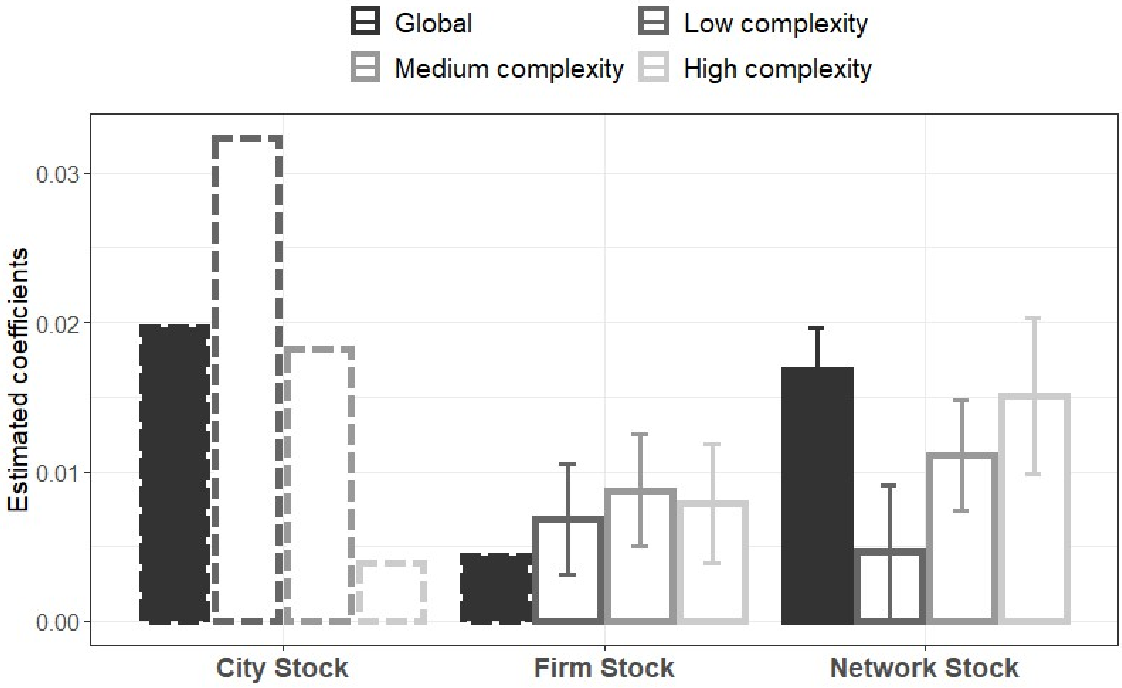
Notes: Marginal effects derived from estimating the main model (equation 1) of the paper, and adding interactions among the three layers considered. Standard errors are clustered at the city-level; the dependent variable is the number of patents per inventor-year, and all variables except "Multinational firm" enter the regressions after an Inverse Hyperbolic Sine (IHS) transformation. Marked x-lines represent the 10th and 90th percentiles, to give an idea of the empirical range of x-axis variable.

Figure 3. Effect of the key variables splitting by knowledge complexity levels



Notes: These coefficients are derived from estimating the main model (equation 1) of the paper by splitting the knowledge stock by complexity levels (one specification for each level of complexity, in order to avoid multicollinearity). Dashed contour indicates a coefficient's p-value <0.05, hence no CI are drawn. Standard errors are clustered at the city-level; the dependent variable is the number of patents per inventor-year, and all variables except "Multinational firm" enter the regressions after an Inverse Hyperbolic Sine (IHS) transformation (see Table A4 in the online Appendix).

Figure 4. Effect of the key variables splitting by knowledge complexity levels. Quality-adjusted patent productivity.



Notes: These coefficients are derived from estimating the main model (equation 1) of the paper by splitting the knowledge stock by complexity levels (one specification for each level of complexity separately, in order to avoid multicollinearity). Dashed contour indicates a coefficient's p-value <0.05, hence no CI are drawn. Standard errors are clustered at the city-level; the dependent variable is the number of patents per inventor-year quality-adjusted, and all variables except "Multinational firm" enter the regressions after an Inverse Hyperbolic Sine (IHS) transformation (see Table A5 in the online Appendix).

Table 1. Top-10 firms by their stock of knowledge in 2010 – with their city location

Firm's name	Firm's city location	Stock of knowledge 2010
PHILLIPS	Eindhoven	3224.41
BASF	Mannheim	2338.96
ROBERT BOSCH	Stuttgart	2136.88
BAYER	Köln	2012.80
SIEMENS	München	1982.28
HOECHST	Frankfurt	1176.25
SIEMENS	Nürnberg	1173.57
BASF	Heidelberg	1104.59
BAYER	Düsseldorf	855.85
CIBA GEIGY	Basel	823.38

Table 2. Baseline results. OLS FE regression with clustered standard errors at the city-level

	(1)	(2)	(3)	(4)
City Stock	0.0482***	0.0444***	0.0409**	0.047***
	(0.0134)	(0.0131)	(0.013)	(0.0132)
Firm Stock		0.0037	0.0014	0.008***
		(0.002)	(0.0018)	(0.0017)
Network Stock			0.0068**	0.0018
			(0.0022)	(0.0019)
City Avg. Productivity				-0.1946
				(0.127)
Firm Avg. Productivity				-0.1271***
				(0.0113)
Network Avg. Productivity				0.0376***
				(0.0066)
Multinational Firm				0.0379***
				(0.0041)
Observations	818,883	818,883	818,883	818,845
Fixed effects				
Inventor	*YES*	*YES*	*YES*	*YES*
Firm	*YES*	*YES*	*YES*	*YES*
City	*YES*	*YES*	*YES*	*YES*
Year	*YES*	*YES*	*YES*	*YES*
Tech	*YES*	*YES*	*YES*	*YES*
Year Tech*	*YES*	*YES*	*YES*	*YES*
City Tech*	*YES*	*YES*	*YES*	*YES*
R^2	0.54515	0.54517	0.54539	0.54628

*Notes: Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors clustered at the city-level. The dependent variable is the number of patents per inventor-year. All variables except "Multinational firm" enter the regressions after an Inverse Hyperbolic Sine (IHS) transformation.*

In knowledge we trust: learning-by-interacting and the productivity of inventors

Online Appendix

Appendix 1: Construction of key variables

Stocks of knowledge computation at the three levels. We compute the stock of knowledge as the cumulative sum of patent applications attributed to a level. A yearly discount rate of 15% applies.

City

a. In order to identify the patent production of cities, we use EUROSTAT Metropolitan Regions to group NUTS3 codes. We retrieve back those NUTS3 not included in a Metropolitan Region whose patent application production is as high as the highest 20% of the same distribution for NUTS3 regions. The resulting set of Metropolitan Regions is our list of cities.
b. The city's stock of knowledge is computed.

Firm

a. Unit of analysis: the match between the firm's HAN code and the CRIOS disambiguated inventor's city address – with the match taking place for inventors of the focal firm in the focal year. The two information altogether signal the existence of a localized firm's R&D establishment.
b. The so-defined firm's stock of knowledge is computed.

Network

a. Each inventor's stock of knowledge is computed.
b. The sum of the stocks belonging to inventors in the focal inventor's network is the network stock. We set a fixed time-window to consider a past collaboration part of the network: from $t - 1$ to $t - 5$. It means that collaborations in the focal year are not considered.

Unique assignment disambiguation. The three levels must match in a single, unique tuple that will have four (non-unique) identifiers: the inventor's code, the firm code, the city code and a time mark. However, some inventors are found to pair with more than one firm and city in the same year, both because of mobility and errors in the source. We set up an algorithm to uniquely assign these "ambiguous" inventors to a city/firm each year. We combine data from PATENTS-ICRIOS database (2014), OECD REGPAT and HAN (2016) databases.

A unique city is assigned to an inventor in a given year according to the following algorithm:

a. Mobile inventors are identified and endowed with three attributes: the *lifespan* (the amount of years they are active in the database); the *nmoves* (the count of moves between regions); the *moveRatio* (equal to *nmoves / lifespan*).
b. Inventors with a *lifespan* = 0 and *nmoves* > 3, together with those in the rightmost 5% of the *moveRatio* distribution, are dropped.

c. For the remaining "ambiguous" inventors, we turn to their patent history. We rank cities they were active in the time window $t \pm 4$ by frequency. The most frequent is chosen.
d. When two cities are equally frequent in the patenting activity of an inventor in the selected time window, the one with a higher knowledge stock is selected.

Following this step, a univocal firm is assigned to an inventor in a given year according to the following algorithm relying on inventors' patent history:

a. We rank firms the inventor was active in the time window $t \pm 4$ by frequency. The most frequent is chosen if appearing 50% more than the second most frequent.
b. The remaining ambiguous assignations are set according to co-location: among firms seemingly frequent, we elect those co-located with the inventor.
c. If more than one firm co-locates with the inventor, the one with the higher knowledge stock is selected.

Patent-based controls: four covariates are computed with patent data in order to control for confounding factors.

A set of *average productivity* covariates, one for each level.

a. *Inventor*. We start measuring productivity (count of applications over time) at the inventor level for a time window ranging from t to $t - 4$. To keep at a minimum the correlation with the knowledge stock variables, instead of the bare count we use the average number of applications belonging to the upper 50% of the patent applications distribution, according to forward-citations received, normalized by year and technology (Waltman et al. 2011; Waltman and Schreiber 2013; Wohlrabe and Bornmann 2017).
b. *City*. The average inventor productivity of inventors belonging to a city at time t is the city covariate.
c. *Firm*. The average inventor productivity of inventors belonging to a firm between t and $t - 3$ is the firm covariate. The time window is necessary to decrease the number of missing observations, emerging from firms not patenting every year.
d. *Network*. The average inventor productivity of inventors belonging to a network from t – 1 to t – 5 is the network covariate. The definition of the network is the same as for the network knowledge stock.

A dummy variable signalling if a firm is a *multinational* or not. Among the possible definitions, we define a firm multinational if at least one of its inventors declares an address in a nation different from that of the firm.

Complexity. We use the Modular Complexity index (MC) introduced by Fleming and Sorenson (2001, 2004). The construction of the modular complexity index relies on the conceptualization of invention as a search in a technological knowledge landscape. Landscapes are made of components, which in turn are measured by technological classes listed in the patent document. The position in the landscape represents a combination of components, with an associated fitness value. Creativity takes the form of a movement on the landscape until a position with a higher fitness appears. The outcome of the search process depends on one factor: the interdependence among technological components. The concept of interdependence coincides with that of modularity or coupling, that is, when two entities are interdependent, a small change in one component calls for changes in the other component in order to the combination to work correctly. We operationalize such procedure as follow.

c. The *Ease of Recombination* (*EoR*) is computed for each technological class-year of the dataset, being technological classes specified as 4-digit IPCs. The *EoR* is the ratio between the count of classes previously combined with the focal class, and the number of applications referencing to the focal class.
d. We calculate *Modular Complexity* (*MC*) index for each patent application as the count of technological classes of the focal patent divided by the sum of their *EoR*, i.e. the inverse of a weighted average.
e. Finally, we compute the stock of patent applications at each level, according to the same methodology explained above, but split by the *MC* index level. More precisely, we assign to *low* MC the applications belonging to the lower 50% of the *MC* index overall distribution, to *medium* MC those belonging to the upper 50-to-90% of the distribution, and to *high* MC the remaining ones.

Before running through this algorithm, we follow the recommendations by Alstott et al. (2017) and apply a normalization procedure, i.e. a Null Model, to the incidence matrix describing occurrences of technological classes within patent applications. In so doing, we clear the probability that two technological classes recombines from a random component originating from technological class and patent populations sizes. Our implemented algorithm is the following:

a. The incidence matrix I of application IDs and IPC codes over a time window from t to $t-4$ is computed.
b. The co-occurrence matrix CO is created as a cross-product of I.
c. We compute *EoR* over the combination of I and CC.
d. 1000 reshuffling of I are elaborated, followed by as many replications of steps 2 and 3.
e. The empirical *EoR* is normalized subtracting the average simulated *EoR* and dividing by its simulated standard deviation, from step 4.

Quality correction. As an alternative dependent variable, we computed the count of high-quality yearly patent applications signed by an inventor. High-quality applications are those belonging to the upper 50% of the patent applications distribution, according to forward-citations received. Citations are constrained within a time window of 5 years after the priority date of the cited patent and normalized by the technological area and cohort (Waltman et al. 2011; Wohlrabe and Bornmann 2017). Citations data are retrieved from PATENTS-ICRIOS database (2014) and collapsed by families to avoid double-counting. The DOC_DB families definition is used.

Appendix 2: Descriptive statistics

Table A.1. Descriptive statistics.

Statistic	N	Mean	St. Dev.	Min	Max
Patents per inventor-year	818,883	1.203	0.503	0.881	6.704
Quality-adj. patents per inventor-year	818,883	0.59	0.629	0	5
City stock	818,883	8.266	1.326	0	10.751
Firm stock	818,883	3.713	2.563	0	9.521
Network stock	818,883	1.883	1.932	0	9
City stock (low complex.)	818,883	7.284	1.305	0	9.728
City stock (medium complex.)	818,883	7.362	1.374	0	9.861
City stock (high complex.)	818,883	6.575	1.479	0	9.382
Firm stock (low complex.)	818,883	2.425	2.408	0	8.46
Firm stock (medium complex.)	818,883	2.66	2.595	0	8.891
Firm stock (high complex.)	818,883	1.941	2.5	0	8.14
Network stock (low complex.)	818,883	0.846	1.371	0	7
Network stock (medium complex.)	818,883	1.09	1.593	0	9
Network stock (high complex.)	818,883	0.683	1.42	0	9
City avg. productivity	818,882	0.244	0.098	0	0.891
Firm avg. productivity	818,845	0.284	0.266	0	3.886
Network avg. productivity	818,883	0.257	0.393	0	5
Multinational	818,883	0.65	0.477	0	1
Specialization index	694,876	0.242	0.017	0.185	0.475
Population density	694,404	0.507	0.313	0.025	2.12
GVApc	694,866	4.013	0.282	1.356	4.943

Note: All variables are IHS-transformed.

Table A.2. Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.59	1										
3	0.08	0.04	1									
4	0.19	0.13	0.36	1								
5	0.32	0.21	0.17	0.35	1							
6	0.14	0.10	0.36	0.30	0.31	1						
7	0.24	0.23	0.21	0.45	0.44	0.44	1					
8	0.28	0.24	0.14	0.31	0.78	0.30	0.50	1				
9	0.14	0.13	0.18	0.51	0.25	0.20	0.29	0.22	1			
10	0.02	0.03	0.15	0.01	0.05	0.15	0.07	0.06	0.02	1		
11	0.04	0.02	0.39	0.13	0.09	0.30	0.14	0.09	0.05	0.35	1	
12	0.08	0.04	0.68	0.23	0.13	0.29	0.19	0.12	0.13	0.21	0.23	1

Note : 1: Patents per inventor-year; 2: Quality-adj. patents per inventor-year; 3: City stock; 4: Firm stock; 5: Network stock; 6: City avg. productivity; 7: Firm avg. productivity; 8: Network avg. productivity; 9: Multinational; 10: Specialization index; 11: Population density; 12: GVA p.c. All variables are IHS-transformed.

Appendix 3: Interaction effects

Table A.3. Two-way interaction effects between main variables of interest. OLS FE regression with clustered standard errors at the city-level

	(1)
City Stock	0.0491***
	(0.0123)
Firm Stock	0.0115
	(0.0083)
Network Stock	0.0043
	(0.0078)
City Avg. Productivity	-0.2578
	(0.1323)
Firm Avg. Productivity	-0.1177***
	(0.0113)
Network Avg. Productivity	0.0297***
	(0.0054)
Multinational Firm	0.0382***
	(0.0042)
*Network Stock*Firm Stock*	0.0035***
	(7e-04)
City Stock Firm Stock*	-9e-04
	(0.0011)
*Network Stock*City Stock*	-0.0021*
	(9e-04)
Observations	818,883
Fixed effects	
Inventor	*YES*
Firm	*YES*
City	*YES*
Year	*YES*
Tech	*YES*
Year Tech*	*YES*
City Tech*	*YES*
R^2	0.54668

*Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors clustered at the city-level. The dependent variable is the number of patents per inventor-year. All variables except "Multinational firm" enter the regressions after an Inverse Hyperbolic Sine (IHS) transformation.*

Appendix 4: Complexity and high-quality patents

Table A.4. Complexity. OLS FE regression with clustered standard errors at the city-level

	(1)	(2)	(3)
City stock low complex	0.0347**		
	(0.0133)		
Firm stock low complex	0.013***		
	(0.0017)		
Network stock low complex	-0.0061***		
	(0.0017)		
City stock medium complex		0.032**	
		(0.0118)	
Firm stock medium complex		0.017***	
		(0.0016)	
Network stock medium complex		-4e-04	
		(0.002)	
City stock high complex			0.0122
			(0.0078)
Firm stock high complex			0.0165***
			(0.0016)
Network stock high complex			0.0032
			(0.0028)
City Avg Productivity	-0.1807	-0.2195	-0.1759
	(0.135)	(0.1336)	(0.1242)
Firm Avg Productivity	-0.1182***	-0.1229***	-0.1173***
	(0.0107)	(0.0109)	(0.0109)
Network Avg Productivity	0.0537***	0.0441***	0.0388***
	(0.0074)	(0.0066)	(0.0064)
Multinational Firm	0.0371***	0.0359***	0.037***
	(0.0042)	(0.0041)	(0.0041)
Observations	818,845	818,845	818,845
Fixed effects			
Inventor	YES	YES	YES
Firm	YES	YES	YES
City	YES	YES	YES
Year	YES	YES	YES
Tech	YES	YES	YES
Year Tech*	YES	YES	YES
City Tech*	YES	YES	YES
R^2	0.54646	0.5467	0.54663

*Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors clustered at the city-level. The dependent variable is the number of patents per inventor-year. All variables except "Multinational firm" enter the regressions after an Inverse Hyperbolic Sine (IHS) transformation.*

Table A.5. Quality-adjusted patents and complexity. OLS FE regression with clustered standard errors at the city-level

Dependent variable: quality-adjusted applications per Inventor-Year				
	(1)	(2)	(3)	(4)
City Stock	0.0197			
	(0.0179)			
Firm Stock	0.0044			
	(0.0023)			
Network Stock	0.0168***			
	(0.0015)			
City low complex.		0.0323		
		(0.0175)		
Firm low complex.		0.0068***		
		(0.0019)		
Network low complex.		0.0046*		
		(0.0023)		
City medium complex.			0.0182	
			(0.0149)	
Firm medium complex.			0.0088***	
			(0.0019)	
Network medium complex.			0.0111***	
			(0.0019)	
City high complex.				0.0039
				(0.0098)
Firm high complex.				0.0079***
				(0.002)
Network high complex.				0.0151***
				(0.0027)
City Avg Productivity	0.1853	0.1562	0.1345	0.1556
	(0.0948)	(0.0987)	(0.099)	(0.0974)
Firm Avg Productivity	-0.3505***	-0.3465***	-0.3479***	-0.3451***
	(0.0245)	(0.0235)	(0.0235)	(0.023)
Network Avg Prod.	-0.1683***	-0.1153***	-0.1337***	-0.1309***
	(0.0166)	(0.0164)	(0.0151)	(0.0134)
Multinational Firm	0.0279***	0.0273***	0.0269***	0.0275***
	(0.005)	(0.005)	(0.005)	(0.005)
Observations	818,845	818,845	818,845	818,845
Fixed effects				
Inventor	YES	YES	YES	YES
Firm	YES	YES	YES	YES
City	YES	YES	YES	YES
Year	YES	YES	YES	YES
Tech	YES	YES	YES	YES
Year Tech*	YES	YES	YES	YES
City Tech*	YES	YES	YES	YES
R^2	0.5290	0.52864	0.52883	0.52888

*Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors clustered at the city-level. The dependent variable is the number of quality-adjusted patents per inventor-year. All variables except "Multinational firm" enter the regressions after an Inverse Hyperbolic Sine (IHS) transformation.*

Appendix 5: Robustness checks

Table A.6. **OLS FE regression with clustered standard errors at the city-level. Global firm's stocks and CE controls**

	(1)	(2)
City Stock	0.0526***	0.0327*
	(0.0125)	(0.0142)
Global Firm Stock	0.004	
	(0.0026)	
Firm Stock		0.0087***
		(0.0021)
Network Stock	0.002	0.0018
	(0.0019)	(0.0021)
City Avg. Productivity	-0.2101	-0.2078
	(0.1229)	(0.1701)
Firm Avg. Productivity	-0.1911***	-0.1304***
	(0.0153)	(0.0125)
Network Avg. Productivity	0.0382***	0.0404***
	(0.0065)	(0.0068)
Multinational Firm	0.0385***	0.0371***
	(0.0042)	(0.0045)
Herfindhal Index		0.3143
		(0.2499)
Pop density		-0.1931
		(0.2927)
GVA per capita		-0.0173
		(0.0496)
Observations	818,845	818,845
Fixed effects		
Inventor	YES	YES
Firm	YES	YES
City	YES	YES
Year	YES	YES
Tech	YES	YES
Year Tech*	YES	YES
City Tech*	YES	YES
R^2	0.54632	0.54523

*Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Standard errors clustered at the city-level. The dependent variable is the number of patents per inventor-year. All variables except "Multinational firm" enter the regressions after an Inverse Hyperbolic Sine (IHS) transformation.*

UBIREA

Institut de Recerca en Economia Aplicada Regional i Públic
Research Institute of Applied Economics

WEBSITE: www.ub-irea.com • **CONTACT:** irea@ub.edu

AQR

Grup de Recerca Anàlisi Quantitativa Regional
Regional Quantitative Analysis Research Group

WEBSITE: www.ub.edu/aqr/ • **CONTACT:** aqr@ub.edu

Universitat de Barcelona
Av. Diagonal, 690 • 08034 Barcelona